

Erratum: Coding map for a contractive Markov system

BY Ivan Werner

*Email: ivan_werner@mail.ru**(Received)**Abstract*

An error in the proof of Lemma 2 (ii) in [I. Werner, Math. Proc. Camb. Phil. Soc. 140(2) 333-347 (2006)], which claims the absolute continuity of dynamically defined measures (DDM), is identified. This undermines the assertion of the positivity of a DDM which provides a construction for equilibrium states in [I. Werner, J. Math. Phys. 52 122701 (2011)]. To rectify that, a dynamical generalization $K^*(\Lambda|\phi_0)$ of the Kullback-Leibler divergence is introduced, which, in the case of its finiteness, allows to obtain a lower bound on the norm of the DDM through

$$\|\Phi\| \geq e^{K(\Lambda|\hat{\Phi}) - K^*(\Lambda|\phi_0)}$$

where $\hat{\Phi}$ is the normed Φ and $K(\Lambda|\hat{\Phi})$ is the Kullback-Leibler divergence. It is shown that $K^*(\Lambda|\phi_0)$ is finite in the case when all maps of a contractive Markov system (CMS) are contractions, the probability functions are Dini-continuous and bounded away from zero and Λ is an equilibrium state of the CMS such that $\Lambda \ll \phi_0$. The question whether the DDM is not zero also in the case of the contraction only on average remains open.

MSC 2010: 28A12, 28A35, 28A80, 82B26, 82C99, 37H99, 37A50, 60J05.

Keywords: Dynamically defined measures, Kullback-Leibler divergence, random systems with complete connections, learning models, g -functions, iterated function systems with place-dependent probabilities, contractive Markov systems, equilibrium states.

Contents

| | | |
|---|--------------------|---|
| 1 | Introduction | 1 |
| 2 | General method | 2 |
| 3 | Application to CMS | 5 |
| | References | 9 |

1. Introduction

Recently, the author has found an error in the last step of the proof of Lemma 2 (ii) in [5] (the author is grateful to Boris M. Gurevich for the invitation to give a talk at

the Ergodic Theory and Statistical Mechanics Seminar at the Lomonosov Moscow State University during the preparation to which the error was discovered). However, the main result of [5] (Corollary 1) is correct even under much weaker conditions on a contractive Markov system (CMS), see Theorem 5 (ii) in [8], than it was required in all articles which used the result (openness of the Markov partition, boundedness away from zero and Dini-continuity of the probability functions). In that respect, the main result of [5] is already obsolete. However, the constructive approach which was taken there gave rise to the technique of dynamically defined measures (DDMs) [6], [7], [9], [10], which allows to construct equilibrium states for such random dynamical systems, and Lemma 2 (ii) in [5] has been the only justification so far that the constructed measure is not zero. In that respect, the article deserves some further consideration.

The lemma was needed for Corollary 1 in [5] only in the special case when ν is substituted with an invariant probability measure μ (we will use the notation from [5]), so that $\Phi(\mu)$ becomes a shift-invariant Borel probability measure M (see Proposition 1 in [5]), $\lambda = 1/N \sum_{i=1}^N \delta_{x_i}$ and all probability functions $p_e|_{K_{i(e)}}$ are Dini-continuous and bounded away from zero. In this case, it is quite easy to see, from the last inequality on page 343 in [5], that even a stronger result is true if all $w_e|_{K_{i(e)}}$'s are contractions on a bounded space, namely there exists $0 < c < \infty$ such that

$$M \leq c\Phi(\lambda).$$

The same holds true if the boundedness condition on the space is removed, but it requires more elaboration (see Corollary 1).

In this note, we first develop a general method for a computation of a lower bound for the norm of a DDM via a dynamical generalization of the *Kullback-Leibler divergence*. (The method can be formulated for an arbitrary invertible dynamics, as also the construction of a DDM, see [10], where the result has already been strengthened by replacing the dynamical generalization of the Kullback-Leibler divergence with the *relative entropy measure*, but it has only a theoretical value so far.) Then we apply the method for a computation of an explicit lower bound for the norm of a DDM associated with a CMS in the case when all its maps are contractions and the probability functions are Dini-continuous and bounded away from zero. No openness of the Markov partition is required, we only assume that M is an equilibrium state (which is, for example, automatically the case when the CMS is non-degenerate [8]). In the case of contraction only on average, it is clear that the DDM is not zero if the probability functions are constant, but there is no proof for that even if the probability functions are only Lipschitz.

2. General method

Now, we will specify our method for the computation of a lower bound for $\Phi(\lambda)(\Sigma)$.

Let ϕ_0 be a probability measure on \mathcal{A}_0 and Φ be the dynamically defined outer measure resulting from ϕ_0 (as in Definition 2 in [5]). Let Λ be a S -invariant Borel probability on Σ . We will slightly abuse the notation by denoting the restriction of Λ on \mathcal{A}_0 also by Λ . Let Z be a version of the Radon-Nikodym derivative $d\Lambda/d\phi_0$. Let $\mathcal{B}(\Sigma)$ denote the Borel σ -algebra on Σ .

We will use the result from [6] that the restriction of Φ on $\mathcal{B}(\Sigma)$ is a S -invariant measure

and the following easily verifiable fact (e.g. Lemma 4 in [10]). For $Q \in \mathcal{B}(\Sigma)$, let $\dot{\mathcal{C}}(Q)$ denote the set of all $(A_m)_{m \leq 0} \in \mathcal{C}(Q)$ such that $A_i \cap A_j = \emptyset$ for all $i \neq j \leq 0$. Then

$$\Phi(Q) = \inf_{(A_m)_{m \leq 0} \in \dot{\mathcal{C}}(Q)} \sum_{m \leq 0} \phi_0(S^m A_m) \quad \text{for all } Q \in \mathcal{B}(\Sigma).$$

Definition 1 Suppose $\Lambda \ll \phi_0$. Set

$$K(\Lambda|\phi_0) := \int \log Z d\Lambda.$$

It is called the *Kullback-Leibler divergence* of Λ with respect to ϕ_0 . Now, define

$$Z^* := \sup_{m \leq 0} Z \circ S^m \quad \text{and}$$

$$K^*(\Lambda|\phi_0) := \int \log Z^* d\Lambda.$$

It is easy to see that $\int \log Z^* d\Lambda$ is well defined, $K(\Lambda|\phi_0) \leq K^*(\Lambda|\phi_0)$ and $K(\Lambda|\phi_0) = K^*(\Lambda|\phi_0)$ if ϕ_0 is S -invariant. Recall that $K(\Lambda|\phi_0) \geq 0$ (which follows immediately from the fact that $x \log x \geq x - 1$ for all $x \geq 0$).

We will use the finiteness of $K^*(\Lambda|\phi_0)$ for a computation of a lower bound for $\Phi(\Sigma)$ through the following lemma (we could refer to a stronger result which already has been obtained in Theorem 5 in [10], but, in this case, a simple proof can be given).

Lemma 1 Suppose $\Lambda \ll \phi_0$ and $K^*(\Lambda|\phi_0) < \infty$. Then

(i)

$$\Phi(Q) \geq \Lambda(Q) e^{-\frac{1}{\Lambda(Q)} \int_Q \log Z^* d\Lambda} \quad \text{for all } Q \in \mathcal{B}(\Sigma) \text{ such that } \Lambda(Q) > 0, \text{ and}$$

(ii)

$$\Phi(\Sigma) \geq e^{K(\Lambda|\hat{\Phi}) - K^*(\Lambda|\phi_0)}$$

where $\hat{\Phi} = \Phi/\Phi(\Sigma)$ (in particular, $K(\Lambda|\hat{\Phi}) \leq K^*(\Lambda|\phi_0)$).

Proof. (i) Note that by the hypothesis $\int \log^+ Z^* d\Lambda < \infty$. Let $Q \in \mathcal{B}(\Sigma)$ such that $\Lambda(Q) > 0$ and $(A_m)_{m \leq 0} \in \dot{\mathcal{C}}(Q)$. Let $0 < \alpha < 1$. Observe that, by the convexity of $x \mapsto e^{-x}$,

$$\begin{aligned} \sum_{m \leq 0} \int_{S^m A_m} Z^{1-\alpha} d\phi_0 &= \sum_{m \leq 0} \int_{S^m A_m} e^{-\alpha \log Z} d\Lambda \\ &\geq \sum_{m \leq 0, \Lambda(A_m) > 0} \Lambda(A_m) e^{-\frac{\alpha}{\Lambda(A_m)} \int_{S^m A_m} \log Z d\Lambda} \\ &\geq \sum_{m \leq 0} \Lambda(A_m) e^{-\frac{\alpha}{\sum_{m \leq 0} \Lambda(A_m)} \sum_{m \leq 0} \int_{S^m A_m} \log Z d\Lambda} \\ &\geq \Lambda(Q) e^{-\frac{\alpha}{\Lambda(\bigcup_{m \leq 0} A_m)} \int_{\bigcup_{m \leq 0} A_m} \log Z^* d\Lambda} \\ &\geq \Lambda(Q) e^{-\alpha \int_Q \log Z^* d\Lambda}. \end{aligned}$$

On the other hand, by the concavity of $x \mapsto x^{1-\alpha}$ or the Hölder inequality, and then by the concavity of $x \mapsto x^\alpha$,

$$\begin{aligned}
\sum_{m \leq 0} \int_{S^m A_m} Z^{1-\alpha} d\phi_0 &\leq \sum_{m \leq 0} \phi_0(S^m A_m)^\alpha \left(\int_{S^m A_m} Z d\phi_0 \right)^{1-\alpha} \\
&= \sum_{m \leq 0} \Lambda(A_m)^{1-\alpha} \phi_0(S^m A_m)^\alpha \\
&\leq \left(\sum_{m \leq 0} \Lambda(A_m) \right) \sum_{m \leq 0} \frac{\Lambda(A_m)}{\sum_{m \leq 0} \Lambda(A_m)} \left(\frac{\phi_0(S^m A_m)}{\Lambda(A_m)} \right)^\alpha \\
&\leq \left(\sum_{m \leq 0} \Lambda(A_m) \right)^{1-\alpha} \left(\sum_{m \leq 0} \phi_0(S^m A_m) \right)^\alpha.
\end{aligned}$$

Hence,

$$\left(\sum_{m \leq 0} \Lambda(A_m) \right)^{1-\alpha} \left(\sum_{m \leq 0} \phi_0(S^m A_m) \right)^\alpha \geq \Lambda(Q) e^{-\frac{\alpha}{\Lambda(\bigcup_{m \leq 0} A_m)} \int_{\bigcup_{m \leq 0} A_m} \log Z^* d\Lambda}.$$

Taking the limit as $\alpha \rightarrow 1$, gives

$$\sum_{m \leq 0} \phi_0(S^m A_m) \geq \Lambda(Q) e^{-\frac{1}{\Lambda(\bigcup_{m \leq 0} A_m)} \int_{\bigcup_{m \leq 0} A_m} \log Z^* d\Lambda}. \quad (2.1)$$

Therefore,

$$\Phi(Q) \geq \Lambda(Q) e^{-\frac{1}{\Lambda(Q)} \int \log^+ Z^* d\Lambda}.$$

This implies indirectly, that $\Lambda \ll \Phi$.

Now, let $(A_m^n)_{m \leq 0} \in \dot{\mathcal{C}}(Q)$ for all $n \in \mathbb{N}$ such that $\sum_{m \leq 0} \phi_m(A_m^n) \downarrow \Phi(Q)$ as $n \rightarrow \infty$. Then, since $\sum_{m \leq 0} \phi_m(A_m^n) \geq \Phi(\bigcup_{m \leq 0} A_m^n) \geq \Phi(Q)$, $\Phi(\bigcup_{m \leq 0} A_m^n \setminus Q) \rightarrow 0$, and therefore, $\Lambda(\bigcup_{m \leq 0} A_m^n \setminus Q) \rightarrow 0$. Hence, by splitting $\log = \log^+ - \log^-$, the finiteness of $K^*(\Lambda|\phi_0)$ implies that

$$\int_{\bigcup_{m \leq 0} A_m^n} \log Z^* d\Lambda \rightarrow \int_Q \log Z^* d\Lambda \quad (\text{as } n \rightarrow \infty).$$

Therefore, by 2.1,

$$\Phi(Q) \geq \Lambda(Q) e^{-\frac{1}{\Lambda(Q)} \int_Q \log Z^* d\Lambda}.$$

This proves (i).

(ii) By (i), for every Borel measurable partition $(Q_k)_{1 \leq k \leq n}$ of Σ ,

$$\sum_{k=1}^n \Lambda(Q_k) \log \frac{\Lambda(Q_k)}{\hat{\Phi}(Q_k)} - \int \log Z^* d\Lambda \leq \log \Phi(\Sigma).$$

Now, using the well-known fact that the sum in the last inequality converges to $K(\Lambda|\hat{\Phi})$ if one chooses a sequence of partitions which is increasing with respect to the refinement

and generates the σ -algebra (e.g. Theorem 4.1 in [3]), it follows that

$$K(\Lambda|\hat{\Phi}) - \int \log Z^* d\Lambda \leq \log \Phi(\Sigma),$$

which proves (ii). \square

3. Application to CMS

We will proceed now towards the application of Theorem 1 for the CMS. Let us abbreviate $\mathcal{M} := (K_{i(e)}, w_e, p_e)_{e \in E}$.

Definition 2 Let $x_i \in K_i$ be fixed for all $i \in N$, as in [5]. Let $\{i \in N : F(M)(K_i) > 0\} \subset S \subset \{1, \dots, N\}$ where $F(M)$ is the measure given by $M \circ F^{-1}$. Set $\lambda' := 1/|S| \sum_{i \in S} \delta_{x_i}$ where $|S|$ denotes the size of S (so that λ from [5] becomes a particular case of λ'). For $m \leq n \in \mathbb{Z}$, let \mathcal{B}_{mn} denote the sub- σ -algebra of \mathcal{A}_m generated by cylinder sets of the form ${}_m[e_m, \dots, e_n]$, $e_i \in E$ for all $m \leq i \leq n$. Let $\Phi_m(\lambda')|_{\mathcal{B}_{mn}}$ denote the restrictions of the measure on the sub- σ -algebra. Then one easily sees by the definitions of the measures that $M|_{\mathcal{B}_{mn}} \ll \Phi_m(\lambda')|_{\mathcal{B}_{mn}}$. Define the Radon-Nikodym derivatives

$$Z_{mn} := \frac{M|_{\mathcal{B}_{mn}}}{\Phi_m(\lambda')|_{\mathcal{B}_{mn}}} \text{ and } Z_{mn}^x := \frac{P_x^m|_{\mathcal{B}_{mn}}}{\Phi_m(\lambda')|_{\mathcal{B}_{mn}}} \text{ for } x \in K.$$

Observe that, since M is S -invariant, and $\Phi_m(\lambda') = \Phi_0(\lambda') \circ S^m$ for all $m \leq 0$, $Z_{m(m+k)} = Z_{0k} \circ S^m$ for all $m \leq 0$ and $k \geq 0$. Furthermore, note that $(Z_{0n}, \mathcal{B}_{0n})_{n \in \mathbb{N}}$ is a $\Phi_0(\lambda')$ -martingale with $\int Z_{0n} d\Phi_0(\lambda') = 1$ for all $n \in \mathbb{N}$. Hence, by Doob's Martingale Theorem, $Z_{0\infty} := \lim_{n \rightarrow \infty} Z_{0n}$ exists $\Phi_0(\lambda')$ -a.e.

Definition 3 Set

$$K_n(M|\Phi_0(\lambda')) := \int Z_{0n} \log Z_{0n} d\Phi_0(\lambda')$$

with the usual continuous extension $0 \log 0 := 0$. It is well known that $0 \leq K_n(M|\Phi_0(\lambda')) \leq K_{n+1}(M|\Phi_0(\lambda'))$ for all $n \in \mathbb{N}$. Set

$$K(M|\Phi_0(\lambda')) := \lim_{n \rightarrow \infty} K_n(M|\Phi_0(\lambda')),$$

which is another way to define the *Kullback-Leibler divergence* of measures.

Recall that $K(M|\Phi_0(\lambda')) < \infty$ implies that $M \ll \Phi_0(\lambda')$ (e.g. Example 4.5.10 in [1] Vol. 1). Furthermore, $M \ll \Phi_0(\lambda')$ implies that $Z_{0\infty} = Z$ $\Phi_0(\lambda')$ -a.e. where Z denotes, from now on, a measurable version of $dM/d\Phi_0(\lambda')$, and $K(M|\Phi_0(\lambda')) = \int Z_{0\infty} \log Z_{0\infty} d\Phi_1(\lambda') = \int \log Z_{0\infty} dM$.

Now, let Σ_G be the set of all two-sided infinite paths of the CMS, i.e.

$$\Sigma_G = \{\sigma \in \Sigma \mid i(\sigma_{i+1}) = t(\sigma_i) \text{ for all } i \in \mathbb{Z}\}.$$

Set

$$b := \sup_{1 \leq i \leq N} \sup_{x \in K_i} \sum_{e \in E, i(e)=i} p_e(x) d(w_e(x_{i(e)}), x_{t(e)})$$

and

$$D := \left\{ \sigma \in \Sigma_G \mid \lim_{m \rightarrow -\infty} w_{\sigma_0} \circ \dots \circ w_{\sigma_m}(x_{i(\sigma_m)}) \text{ exists} \right\}.$$

For $\sigma \in \Sigma$, let

$$F(\sigma) := \begin{cases} \lim_{m \rightarrow -\infty} w_{\sigma_0} \circ w_{\sigma_{-1}} \circ \dots \circ w_{\sigma_m}(x_{i(\sigma_m)}) & \text{if } \sigma \in D \\ x_{t(\sigma_0)} & \text{otherwise.} \end{cases}$$

Let \mathcal{F} denote the σ -algebra generated by cylinder sets of the form $m[e_m, \dots, e_0]$, $e_m, \dots, e_0 \in E$, $m \leq 0$. Set

$$E(\mathcal{M}) := \{ \Lambda \in P_S(\Sigma) \mid \Lambda(D) = 1 \text{ and } E_\Lambda(1_{1[e]}|\mathcal{F}) = p_e \circ F \text{ } \Lambda\text{-a.e. for all } e \in E \}.$$

By Corollary 1 (ii) in [8], $M \in E(\mathcal{M})$ if \mathcal{M} is contractive, uniformly continuous and non-degenerate (see [8] for the definition of the non-degeneracy, \mathcal{M} is non-degenerate, in particular, if all K_i ' are open). This already contains the main result from [5] that $M(D) = 1$ (by Theorem 5 (ii) in [8], $M(D) = 1$ even under much weaker conditions), but we are now concerned with a proof that the outer measure $\Phi(\lambda')$ constructed in [5], which was shown to define a S -invariant Borel measure in [6], is not zero.

Remark 1 Note that for every $\Lambda \in E(\mathcal{M})$, $1_{K_{i(\sigma_1)}} \circ F(\sigma) = 1$ for Λ -a.e. $\sigma \in \Sigma$, as

$$\begin{aligned} \int 1_{K_{i(\sigma_1)}} \circ F(\sigma) d\Lambda(\sigma) &= \sum_{j=1}^N \sum_{e \in E, i(e)=j} \int 1_{1[e]} 1_{K_j} \circ F d\Lambda = \sum_{j=1}^N \sum_{e \in E, i(e)=j} \int p_e \circ F 1_{K_j} \circ F d\Lambda \\ &= \sum_{j=1}^N \int 1_{K_j} \circ F d\Lambda = 1. \end{aligned}$$

Therefore, for continuous $w_e|_{K_{i(e)}}$'s,

$$w_{\sigma_0} \circ \dots \circ w_{\sigma_{m+1}} \circ F(S^m \sigma) = F(\sigma) \text{ for } \Lambda\text{-a.e. } \sigma \in \Sigma \text{ and all } m \leq 0.$$

Set

$$\begin{aligned} C &:= \sum_{j=1}^N \int_{K_j} d(x, x_j) d\mu(x), \\ \Delta(t) &:= \sup_{e \in E} \sup_{x, y \in K_{i(e)}, d(x, y) \leq t} |p_e(x) - p_e(y)| \text{ for all } t \geq 0, \text{ and} \\ d &:= \sup_{e \in E} d(w_e x_{i(e)}, x_{t(e)}). \end{aligned}$$

Note that, by Lemma 14 in [8], $C < b/(1-a)$ (clearly, $b \leq d < \infty$, since E is finite). In the case when all $w_e|_{K_{i(e)}}$ are contraction, this can be strengthened to the following.

Lemma 2 *If all $w_e|_{K_{i(e)}}$ are contraction with a contraction rate $0 < a < 1$, then*

$$d(F(\sigma), x_{i(\sigma_1)}) \leq \frac{d}{1-a} \text{ for all } \sigma \in \Sigma_G.$$

Proof. Let $\sigma \in \Sigma_G$. Let us abbreviate

$$X_m(\sigma) := w_{\sigma_0} \circ w_{\sigma_{-1}} \circ \dots \circ w_{\sigma_m}(x_{i(\sigma_m)}) \text{ for all } m \leq 0 \text{ and } X_1(\sigma) := x_{i(\sigma_1)}.$$

Then, for $m \leq 0$,

$$d(X_m(\sigma), x_{i(\sigma_1)}) \leq \sum_{m \leq k \leq 0} d(X_k(\sigma), X_{k+1}(\sigma)) \leq \sum_{m \leq k \leq 0} a^{-k} d.$$

Thus taking the limit, as $m \rightarrow -\infty$, implies the assertion. \square

Theorem 1 Suppose \mathcal{M} is contractive with a contraction rate $0 < a < 1$, $p_e|_{K_{i(e)}}$'s are Dini-continuous, and there exists $\delta > 0$ such that $p_e|_{K_{i(e)}} \geq \delta$ for all $e \in E$. Suppose $M \in E(\mathcal{M})$. Then the following holds true.

(i)

$$K(M|\Phi_0(\lambda')) \leq \log |S| + \frac{1}{\delta} \left[\frac{1}{1 - \sqrt{a}} + \sum_{i=0}^{\infty} \Delta \left(a^{\frac{i}{2}} C \right) \right].$$

(ii) If all $w_e|_{K_{i(e)}}$ are contractions with a contraction rate $0 < a < 1$, then

$$\log Z_{0\infty} \leq \log |S| + \frac{1}{\delta} \sum_{i=0}^{\infty} \Delta \left(a^i \frac{d}{1-a} \right) \quad M\text{-a.e.}$$

Proof. Observe that the hypothesis implies that $w_e|_{K_{i(e)}}$'s are Lipschitz. Also, note that $Z_{0\infty} = Z_{1\infty} \circ S^{-1}$ M -a.e.

(i) Let $n \in \mathbb{N}$, and (e_1, \dots, e_n) be a path. Observe that, by the convexity of $t \mapsto t \log t$,

$$\begin{aligned} & M(1[e_1, \dots, e_n]) \log \frac{M(1[e_1, \dots, e_n])}{\frac{1}{|S|} P_{x_{i(e_1)}}^1(1[e_1, \dots, e_n])} \\ &= \int P_x^1(1[e_1, \dots, e_n]) d\mu(x) \log \frac{\int P_x^1(1[e_1, \dots, e_n]) d\mu(x)}{\frac{1}{|S|} P_{x_{i(e_1)}}^1(1[e_1, \dots, e_n])} \\ &\leq \int P_x^1(1[e_1, \dots, e_n]) \log \frac{P_x^1(1[e_1, \dots, e_n])}{\frac{1}{|S|} P_{x_{i(e_1)}}^1(1[e_1, \dots, e_n])} d\mu(x). \end{aligned}$$

Hence,

$$\int_{1[e_1, \dots, e_n]} \log Z_{1n} dM \leq \int_{1[e_1, \dots, e_n]} \int \log Z_{1n}^x dP_x^1 d\mu(x)$$

for all $1[e_1, \dots, e_n] \in \mathcal{B}_{1n}$, and therefore,

$$\int_A \log Z_{1n} dM \leq \int_A \int \log Z_{1n}^x dP_x^1 d\mu(x) \quad (3.1)$$

for all $A \in \mathcal{B}_{1n}$. Let $A \in \mathcal{B}_{1n}$. By Proposition 1 in [5] and Theorem 5 (ii) in [8] (Theorem 3.27 (ii) in the DSDC version), $\mu = F(M)$. Then, since $M \in E(\mathcal{M})$, by Lemma 4 (ii)

(Lemma 3.8 (ii) in the DSDC version) in [8],

$$\begin{aligned}
\int_A \log Z_{1n} dM &\leq \int_A \log Z_{1n}^{F(\sigma)}(\sigma) dM(\sigma) \\
&= \int_A \log \frac{p_{\sigma_1}(F(\sigma)) \dots p_{\sigma_n}(w_{\sigma_{n-1}} \circ \dots \circ w_{\sigma_1} \circ F(\sigma))}{\frac{1}{|S|} p_{\sigma_1}(x_{i(\sigma_1)}) \dots p_{\sigma_n}(w_{\sigma_{n-1}} \circ \dots \circ w_{\sigma_1} x_{i(\sigma_1)})} dM(\sigma) \\
&= M(A) \log |S| + \sum_{i=1}^n \int_A \log \frac{p_{\sigma_i}(w_{\sigma_{i-1}} \circ \dots \circ w_{\sigma_1} \circ F(\sigma))}{p_{\sigma_i}(w_{\sigma_{i-1}} \circ \dots \circ w_{\sigma_1} x_{i(\sigma_1)})} dM(\sigma).
\end{aligned}$$

Using the inequality $\log(x) \leq x - 1$, it follows that

$$\begin{aligned}
&\int_A \log Z_{1n} dM \\
&\leq M(A) \log |S| \\
&\quad + \sum_{i=1}^n \int_A \frac{|p_{\sigma_i}(w_{\sigma_{i-1}} \circ \dots \circ w_{\sigma_1} \circ F(\sigma)) - p_{\sigma_i}(w_{\sigma_{i-1}} \circ \dots \circ w_{\sigma_1} x_{i(\sigma_1)})|}{p_{\sigma_i}(w_{\sigma_{i-1}} \circ \dots \circ w_{\sigma_1} x_{i(\sigma_1)})} dM(\sigma)
\end{aligned}$$

for all $A \in \mathcal{B}_{1n}$. For $\sigma \in \Sigma$, set

$$f(\sigma) := \sum_{i=1}^{\infty} |p_{\sigma_i}(w_{\sigma_{i-1}} \circ \dots \circ w_{\sigma_1} \circ F(\sigma)) - p_{\sigma_i}(w_{\sigma_{i-1}} \circ \dots \circ w_{\sigma_1} x_{i(\sigma_1)})|. \quad (3.2)$$

Then, by the Monotone Convergence Theorem,

$$\int_A \log Z_{1n} dM \leq M(A) \log |S| + \frac{1}{\delta} \int_A f dM \quad (3.3)$$

for all $A \in \mathcal{B}_{1n}$. We show that $f \in \mathcal{L}^1(M)$. Using Lemma 4 (ii) in [8] (Lemma 3.8 (ii) in the DSDC version), and the contraction on average,

$$\begin{aligned}
&\int d(w_{\sigma_i} \circ \dots \circ w_{\sigma_1} \circ F(\sigma), w_{\sigma_i} \circ \dots \circ w_{\sigma_1} x_{i(\sigma_1)}) dM(\sigma) \\
&\leq \int \int d(w_{\sigma_i} \circ \dots \circ w_{\sigma_1} x, w_{\sigma_i} \circ \dots \circ w_{\sigma_1} x_{i(\sigma_1)}) dP_x^1(\sigma) d\mu(x) \\
&\leq a^i C
\end{aligned} \quad (3.4)$$

for all $i \in \mathbb{N}$. For $i \in \mathbb{N}$, set

$$A_i := \left\{ \sigma \in \Sigma \mid d(w_{\sigma_i} \circ \dots \circ w_{\sigma_1} \circ F(\sigma), w_{\sigma_i} \circ \dots \circ w_{\sigma_1} x_{i(\sigma_1)}) > a^{\frac{i}{2}} C \right\}$$

Then, by (3.4),

$$M(A_i) \leq a^{\frac{i}{2}} \text{ for all } i \in \mathbb{N}.$$

Then

$$\int f dM \leq \sum_{i=1}^{\infty} M(A_{i-1}) + \sum_{i=1}^{\infty} \Delta(a^{\frac{i-1}{2}} C) \leq \frac{1}{1-\sqrt{a}} + \sum_{i=1}^{\infty} \Delta(a^{\frac{i-1}{2}} C)$$

for all $n \in \mathbb{N}$. Hence, by the hypothesis, $f \in \mathcal{L}^1(M)$. Furthermore, by (3.3),

$$\log Z_{1n} \leq \log |S| + \frac{1}{\delta} E_M(f | \mathcal{B}_{1n}) \quad M\text{-a.e.}$$

Therefore, since $\mathcal{B}_{1n} \uparrow \mathcal{A}_1$ ($n \rightarrow \infty$),

$$\log Z_{1\infty} \leq \log |S| + \frac{1}{\delta} E_M(f|\mathcal{A}_1) \quad M\text{-a.e.} \quad (3.5)$$

By Shiryaev's Local Absolute Continuity Theorem (e.g. Theorem 2, p. 514 in [2]), this implies that $M \ll \Phi_1(\lambda')$ and $Z_{1\infty} = d\Phi_1(\lambda')/dM$. Thus, since $Z_{0\infty} = Z_{1\infty} \circ S^{-1}$, the integration of (3.5) with respect to M implies the assertion.

(ii) By Remark 1 and Lemma 2, for M -a.e. $\sigma \in \Sigma_G$,

$$\begin{aligned} f(\sigma) &\leq \sum_{i=0}^{\infty} \Delta \left(d(w_{\sigma_i} \circ \dots \circ w_{\sigma_1} \circ F(\sigma), w_{\sigma_i} \circ \dots \circ w_{\sigma_1} x_{i(\sigma_1)}) \right) \\ &\leq \sum_{i=0}^{\infty} \Delta \left(a^i d(F(\sigma), x_{i(\sigma_1)}) \right) \\ &\leq \sum_{i=0}^{\infty} \Delta \left(a^i \frac{d}{1-a} \right). \end{aligned}$$

Hence, by (3.5), since $M(\Sigma_G) = 1$,

$$\log Z_{1\infty} \leq \log |S| + \frac{1}{\delta} \sum_{i=0}^{\infty} \Delta \left(a^i \frac{d}{1-a} \right) \quad M\text{-a.e.}$$

Thus the assertion follows. \square

For the definition of the *non-degeneracy* and sufficient conditions for it, we refer to [8] (the non-degeneracy is satisfied if all K_i 's are open, but it also admits a large class of systems with proper Borel-measurable partitions).

Corollary 1 *Suppose \mathcal{M} is non-degenerate such that all $p_e|_{K_{i(e)}}$'s are Dini-continuous, there exists $\delta > 0$ such that $p_e|_{K_{i(e)}} \geq \delta$ for all $e \in E$ and all $w_e|_{K_{i(e)}}$'s are contractions with a contraction rate $0 < a < 1$. Then*

$$\Phi(\lambda')(Q) \geq M(Q) \frac{1}{|S|} e^{-\frac{1}{\delta} \sum_{i=0}^{\infty} \Delta(a^i \frac{d}{1-a})} \quad \text{for all } Q \in \mathcal{B}(\Sigma).$$

Proof. By Corollary 1 (ii) in [8], $M \in E(\mathcal{M})$. Thus the assertion follows by Lemma 1 (i) and Theorem 1 (ii). \square

REFERENCES

- [1] V. I. Bogachev, *Measure theory*, Vol. I, II, Springer (2007).
- [2] A. Shiryaev, *Probability* (in Russian), Nauka (1989).
- [3] W. Słomczyński, Dynamical entropy, Markov operators, and iterated function systems, *Rozprawy Habilitacyjne Uniwersytetu Jagiellońskiego Nr 362, Wydawnictwo Uniwersytetu Jagiellońskiego*, Kraków (2003).
- [4] I. Werner, Contractive Markov systems, *J. London Math. Soc.* **71** (2005) 236-258.
- [5] I. Werner, Coding map for a contractive Markov system, *Math. Proc. Camb. Phil. Soc.* **140** (2) (2006) 333-347, arXiv:math/0504247.
- [6] I. Werner, Dynamically defined measures and equilibrium states, *J. Math. Phys.* **52** (2011) 122701, arXiv:1101.2623.
- [7] I. Werner, Erratum: Dynamically defined measures and equilibrium states, *J. Math. Phys.* **53** 079902 (2012), arXiv:1101.2623.

- [8] I. Werner, Equilibrium states and invariant measures for random dynamical systems, *DCDS-A* **35** (3) (March 2015) 1285 - 1326, arXiv:1203.6432.
- [9] I. Werner, Erratum II: Dynamically defined measures and equilibrium states, arXiv:1101.2623.
- [10] I. Werner, Lower bounds for the dynamically defined measures, arXiv:1506.04497.